PLISKA

STUDIA MATHEMATICA BULGARICA

ПЛИСКА

БЪЛГАРСКИ МАТЕМАТИЧЕСКИ СТУДИИ

Pliska Stud. Math. Bulgar. **22** (2013), 169–194

PLISKA
STUDIA MATHEMATICA
BULGARICA

TIME TO EXTINCTION IN BRANCHING PROCESSES AND ITS APPLICATION IN EPIDEMIOLOGY*

M. Slavtchova–Bojkova

Abstract. The contemporary state of the theory of branching processes implies their application to any abstract population where individuals produce a set of new individuals. In this survey paper some recent developments in the study of time to extinction of continuous–time branching processes (BP) motivated by their applications in epidemiological modeling will be presented. The developed methodology and results are concerning Bellman–Harris (age-dependent) BP and more general Sevast'yanov's BP, as well.

1. Introduction. The Bellman–Harris branching process (BHBP) is a continuous-time model, which has been widely studied in the stochastic processes theory (see for example Chapter IV in Ahtreya et al. (1972)). Moreover, from a practical outlook, it has been used to describe the evolution of populations along time in different situations, as for example, to solve many problems related to cell populations (see Axelrod et al. (1993, 1997), Kimmel (1985), Kimmel et al. (1986), Yakovlev et al. (2006, 2007) and others).

It is well–known that a BHBP becomes extinct or explode to infinity depending on the mean value of its reproduction law. This property is inherited from its embedding Bienaymé–Galton–Watson branching process (EBGWBP), leading us

*This work is partially supported by the financial funds allocated to the Sofia University "St. Kl. Ohridski", grant No 125/2012.

2010 *Mathematics Subject Classification:* Primary 60J80; Secondary 92D30.

Key words: age–dependent branching process, extinction time, epidemiological modeling, vaccination level.

to the classification of subcritical, critical and supercritical cases. Then, the extinction happens almost surely (a.s.) in the subcritical and critical cases, and has a positive probability in the supercritical case (obviously under the corresponding conditions to avoid trivial cases).

However, the time necessary for the extinction of a BHBP can not be deduced from its EBGWBP. This time is a random variable (r.v.) which depends on the continuous–time structure of the BHBP on its own. Even though the study of the extinction time is very interesting from both theoretical and practical view points, it has not been considered deeply enough (see Agresti (1974), Farrington et al. (1999), Hainzmann (2009), Pakes (1989)). Gonzáles et al. (2010a) deal with this problem, investigating the dependence of the extinction time of a BHBP on its reproduction law. Moreover, they apply the obtained results in an epidemiological context. Actually, the problem of how to model the evolution of an infectious disease is very important and widely considered in the recent literature (see Becker et al. (2004), Farrington et al. (2003b), Isham (2005), Mode et al. (2000) and Pakes (1989)). However only in few papers (see Andersson et al. (2000), Barbour (1975), Farrington et al. (1999) and Nasell (2002)) the waiting time to extinction of the disease has been used as a main tool to determine a vaccination policy. Mainly because there are not enough results on this r.v. In the work of Gonzáles et al. (2010a) a new approach to this topic was suggested. In Section 2 some properties of the distribution function (d.f.) of the extinction time of a BHBP, mainly those related to stochastic monotonicity and continuity depending on its reproduction law, were studied. Then, this study is applied to investigate the behavior of the time until an infectious disease become extinct depending on the proportion of the immune individuals in the population. The diseases under consideration are those which follow a SIR (susceptible–infected–removed) scheme. It is well–known that branching processes fit adequately this scheme (see Andersson et al. (2000) and Ball et al. (1995)). So, first, the spread of infection is modelled by a BHBP. Then its extinction time distribution was studied and an optimal vaccination level to immunize susceptible individuals in the population is proposed.

In Section 3 the number of infectious individuals in the population depending on the vaccination level was modeled by means of Sevast'yanov's age–dependent branching processes (SBP) (see Sevast'yanov (1971)). This model is a particular case of the general branching process (see Jagers (1975)), also called Crump–Mode–Jagers branching process (CMJBP), which is the most adequate model to fit infectious diseases following SIR scheme (see Ball et al. (1995)). The SBP is specially adequate to model the evolution of diseases with incubation period

(and a negligible contact period) for which the virulence of the disease could be a function of this period. Therefore, using SBPs, our target is to determine the optimal proportion of susceptible individuals which might be vaccinated to guarantee the extinction of the disease within a given period of time. An advance of results without proofs has been published in Gonzáles et al. (2009).

In Section 3.1 the spread of the disease is modeled by SBPs which depend on the proportion of immune individuals in the population. For that reason the time to extinction of an infectious disease is considered, depending on the proportion of immune individuals into the population. Then, the main monotonicity and continuity properties of the time to extinction is studied. A policy for defining the optimal vaccination level, based on the mean of the extinction time distribution of the disease is also suggested. At the end of this section the data from avian influenza spreading in Vietnam at the end of 2006 is analyzed.

In an attempt to meet the threats of infectious diseases to society, public health authorities have created comprehensive mechanisms for the collection of disease data. As a consequence, the abundance of data has demanded the development of automated algorithms for the detection of abnormalities and aberrations. Typically, such an algorithm monitors a univariate time series of counts using a combination of heuristic methods and statistical modeling. Prominent examples of surveillance algorithms are the works by Stroup et al. (1989) and Farrington et al. (1996). A comprehensive survey of outbreak detection methods can be found in Farrington et al. (2003a). The R–package surveillance was written with the aim of providing a test-bench for surveillance algorithms. The purpose of the analysis in Mitova-Bobcheva et al. (2011) is to illustrate the basic functionality of the package with R–code examples.

Section 4 contains a short description of how to use the surveillance algorithms and presents the results with description of the data set used. In Section 5 an overview of Bayesian estimation of the offspring mean of a BGWBP, identified as a basic reproduction number in epidemiology, is presented. Finally this approach is illustrated on the mumps data collected in Bulgaria during the period 2005–2008. The method relies on the BGWBP as a model of epidemic spread. That is why it is reasonable here to remind shortly its well–known definition and to clarify its direct interpretation in epidemiological context.

A more detailed exposition of the theory of branching processes can be found, for example in Jagers (1975) or Slavtchova–Bojkova et al. (2007).

Let $X_i(n)$ are independent and identically distributed random variables (i.i.d. r.v.) with the same distribution as X. The distribution of X is called offspring distribution, the mean of X is denoted by $\lambda = EX$. Formally, we define $\{Z_n, n =$

$0, 1, 2, \ldots\}$ as follows:

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_i(n-1), \quad Z_0 = s,$$

where $X_i(n-1)$ is the number of infected by i-th individual of $(n-1)$-th generation. The sequence of r.v. $\{Z_n, n = 0, 1, 2, \ldots\}$ is clearly a BGWBP.

The interpretation of mathematical model in epidemiological context is clear: it is assumed that each infectious individual infects a random number of susceptible individuals distributed as a r.v. X. Let us start with s infected individuals. All infected individuals due to a contact with them are called first generation, and let us denote their number by Z_1. Infected individuals in contact with the first generation form the second generation, with Z_2 individuals, etc.

The event $\{Z_n = 0, \quad \text{for some} \quad n \geq 1 \mid Z_0 = 1\}$ is called extinction. Denote the probability of extinction $q = P\{Z_n = 0, \quad \text{for some} \quad n \geq 1 \mid Z_0 = 1\}$. From the theory of branching processes it is known that for $\lambda \leq 1$, $q = 1$, and for $\lambda > 1$, $q < 1$.

If the process starts with s individuals, the probability of extinction is $P\{Z_n = 0, n \geq 1 \mid Z_0 = s\} = q^s$.

Depending on whether the offspring mean λ is less than, equal to or greater than 1, process is called subcritical, critical and supercritical, respectively.

2. Age-dependent (Bellman–Harris) branching processes.

2.1. Properties of the extinction time of BHBP. In this section some properties related to the extinction time of BHBPs are presented. First we draw our attention on obtaining results concerning a BHBP with fixed reproduction law, which is referenced in terms of its probability generating function (p.g.f.). Then, we study the properties of the extinction time of BHBPs with different reproduction laws but with the same distribution of the life–length. Specifically, we establish stochastic monotonicity and continuity properties depending on the reproduction law. From now on the same notation $\{Z_t\}_{t\geq 0}$ is used for the BP in question.

To this aim, by T_f is denoted the extinction time of a BHBP, $\{Z_t\}_{t\geq 0}$, initiated at time 0 with a single individual, with reproduction law given by its p.g.f. $f(\cdot)$ and life-length with d.f. $G(\cdot)$ such that $G(0^+) = 0$. Mathematically, we have

$$T_f = \inf\{t \geq 0 : Z_t = 0\},$$

where Z_t denotes the number of individuals of the population at time t.

Fixed the p.g.f. $f(\cdot)$, the d.f. of the extinction time T_f is denoted by $v_f(\cdot)$, i.e.

$$v_f(t) = P(T_f \leq t),\ t \in \mathbb{R}.$$

Since $G(0^+) = 0$, then $v_f(0) = 0$. Furthermore, using the methods given in the book by Athreya et al. (1972) (see p. 139, Theorem IV.2.1), it is easy to deduce that $v_f(\cdot)$ is the unique bounded function that satisfies the integral equation:

$$v_f(t) = \begin{cases} 0, & t < 0, \\ \displaystyle\int_0^t f(v_f(t-s))dG(s), & t \geq 0. \end{cases} \tag{1}$$

Moreover, let q_f be the extinction probability of a BHBP started with one ancestor and with reproduction law given by its p.g.f. $f(\cdot)$. It is clear that $q_f = P(T_f < \infty)$ and it is also well-known that $q_f = 1$ iff $m_f \leq 1$, where m_f denotes the reproduction mean associated to $f(\cdot)$. So that, for such a p.g.f. $f(\cdot)$ with $m_f > 1$, $v_f(\cdot)$ is the d.f. of a non-proper r.v. because $P(T_f < \infty) < 1$. In any case, it follows that

$$\widetilde{v}_f(t) = P(T_f \leq t | T_f < \infty) = \frac{v_f(t)}{q_f},\ t \geq 0, \tag{2}$$

and from (1) it is easy to obtain that $\widetilde{v}_f(\cdot)$ also satisfies the equation

$$\widetilde{v}_f(t) = \int_0^t g(\widetilde{v}_f(t-s))dG(s),\ t \geq 0,$$

where $g(s) = q_f^{-1} f(q_f s)$ is a p.g.f. such that $m_g < 1$, that is, $\widetilde{v}_f(t) = v_g(t)$, for all $t \in \mathbb{R}$.

Therefore, without loss of generality, from now on, in many situations a p.g.f. $f(\cdot)$ can be considered so that the extinction time T_f is a proper r.v., i.e. $m_f \leq 1$.

The following results, which later on are giving us the possibility to develop the simulation-based methodology to approximate the $v_f(\cdot)$, are derived by González et al. (2010a).

Proposition 1. *If $G(\cdot)$ is an absolutely continuous d.f., then $v_f(\cdot)$ is also an absolutely continuous d.f.*

To approximate the value of the d.f. $v_f(\cdot)$ on each point t, the functional operator $H_f(\cdot)$, defined on any function $u(\cdot)$ from the non-negative real numbers $\mathbb{R}_+$ to the closed interval $[0, 1]$ was introduced, as follows:

$$H_f(u)(t) = \int_0^t f(u(t-s))dG(s),\ t \geq 0.$$

Also, for all $n \geq 1$, by $H_f^n(\cdot)$ is denoted the n^{th} composition of the operator $H_f(\cdot)$, that is, $H_f^{n+1}(u)(\cdot) = H_f(H_f^n(u))(\cdot)$, $n = 1, 2, \ldots$ and $H_f^1(u)(\cdot) = H_f(u)(\cdot)$. Using this notation, from (1) it is obtained that $v_f(\cdot)$ is the unique bounded function satisfying the fixed-point equation $u(\cdot) = H_f(u)(\cdot)$.

Theorem 1. *If $f(\cdot)$ is a p.g.f., then for each function $h : \mathbb{R}_+ \to [0, 1]$, it is verified that*

$$v_f(t) = \lim_{n\to\infty} H_f^n(h)(t),\ t \geq 0. \tag{3}$$

Theorem 2. *Let $f(\cdot)$ and $g(\cdot)$ be p.g.f. If $f(s) \leq g(s)$ for all $0 \leq s \leq 1$, then $v_f(t) \leq v_g(t)$ for all $t \geq 0$.*

Theorem 3. *Let $f(\cdot)$ be a p.g.f. such that $m_f < 1$. For each $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon, f) > 0$ such that if $g(\cdot)$ is a p.g.f. satisfying*

$$\sup_{0\leq s\leq 1} |f(s) - g(s)| \leq \delta,$$

then

$$\sup_{0\leq t<\infty} |v_f(t) - v_g(t)| \leq \varepsilon.$$

2.2. Application to epidemic modelling. Branching processes have been widely used to describe the evolution of an infectious disease following a SIR scheme, at least in their early stages, (see Andersson et al. (2000), Ball et al. (1995), Haccou et al. (2005), Kimmel et al. (2002), Mode et al. (2000) and Pakes (2003)). In particular, infectious diseases with long incubation period and negligible contagious time, such as avian flu, measles, mumps, can be described by a BHBP.

To model the spread of an infectious disease by BHBP, the following scheme was considered. It is assumed that three types of individuals may exist in the population: infected, susceptible individuals, and healthy and/or immune to this disease. The disease is spreading when an infected individual is in contact with susceptible individuals. Notice that during the incubation period, the infected individual as yet neither shows any symptoms of the disease nor passes the disease to any susceptible individual. Moreover, when the infectious disease is observed

in an individual, this individual is either isolated (for example in human or animal populations) or culled. Hence, just after the incubation period and before to be isolated or culled, there is a very short contact period (in comparison with the incubation one) in which the individual may infects others.

By p_k is denoted the probability that one infected individual contacts k healthy individuals, $k \geq 0$, and by α $(0 \leq \alpha \leq 1)$ the proportion of immune individuals of the population. Both infected and immune individuals are dispersed uniformly in the population. Furthermore, the population size is fixed and large enough in comparison with the number of infected individuals, so that α and the contact distribution law, $\{p_k\}_{k\geq 0}$, can be considered stable along time (see Isham (2005)). Notice that this is neither a restriction in critical and subcritical processes because of their almost sure extinction, nor in the early stages of supercritical processes.

Under these assumptions, the probability that an infected individual transmits the disease to k susceptible individuals when α is the vaccination level in the population, is given by

$$p_{\alpha,k} = \sum_{j=k}^{\infty} \binom{j}{k} \alpha^{j-k}(1-\alpha)^k p_j, \tag{4}$$

i.e. the infected individual has been in contact with j healthy individuals and among them there have been k susceptible individuals. We call $\{p_{\alpha,k}\}_{k\geq 0}$ the infection distribution law when the proportion of immune individuals of the population is α.

Following this spreading scheme along time, infected individuals pass on the disease to other susceptible individuals and so on. The number of infected individuals in a population with vaccination level α is modelled by a BHBP, whose offspring law is determined by the infection distribution law $\{p_{\alpha,k}\}_{k\geq 0}$ and the d.f. of the life-length of an infected individual is given by an arbitrary d.f. $G(\cdot)$ of a non-negative r.v. By life-length we mean the period (measured in real time) till either he/she infects susceptible individuals or the disease disappears in this individual, that is, the incubation period. Notice that the life-length of an infected individual is assumed to depend neither on the proportion of immune individuals nor on the contact distribution law.

In order to vaccinate a proportion of susceptible individuals, it is supposed that a vaccination policy is applied. The objective is to determine what proportion, α, of these individuals might be vaccinated/immunized to guarantee the extinction of the disease, possibly in a given period of time.

2.3. The extinction time of the epidemic. In what follows, the distribution of the extinction time of a BHBP depending on the vaccination level α is investigated. To this end, for each α such that $0 \leq \alpha \leq 1$, by $f_\alpha(\cdot)$ the p.g.f. of $\{p_{\alpha,k}\}_{k\geq 0}$ is denoted. From (4) it is easily obtained that

$$f_\alpha(s) = f(\alpha + (1-\alpha)s),\ 0 \leq s \leq 1, \tag{5}$$

being $f(\cdot)$ the p.g.f. of $\{p_k\}_{k\geq 0}$. Moreover, by T_α is denoted the extinction time of a BHBP initiated at time 0 with a single infected individual and with p.g.f. $f_\alpha(\cdot)$ and by $v_\alpha(\cdot)$ the d.f. of T_α.

The mean of contacts of an infected individual is denoted by m and by m_α the mean of susceptible individuals, who are infected by a contagious individual, given vaccination level α. Then, from (4) it is obtained that

$$m_\alpha = (1-\alpha)m. \tag{6}$$

Taking into account (6), $m_\alpha \leq 1$ is equivalent to $\max\{0, 1 - m^{-1}\} \leq \alpha \leq 1$, which depends on the mean of contacts of an infected individual. From now on by $\alpha_{\inf} = \max\{0, 1 - m^{-1}\}$ is denoted the smallest proportion of immune individuals, so that the infectious disease becomes extinct a.s.

For fixed α and p, with $\alpha_{\inf} \leq \alpha \leq 1$ and $0 < p < 1$, by t_p^α is denoted the quantile of order p of the variable T_α. The following result holds:

Theorem 4. *Let p be such that $0 < p < 1$.*

1. *If $\alpha_{\inf} \leq \alpha_1 < \alpha_2 \leq 1$, then $t_p^{\alpha_2} \leq t_p^{\alpha_1}$.*
2. *If α is such that $0 < m_\alpha < m_{\alpha_{\inf}}$, then $\lim_{\widetilde{\alpha}\to\alpha^+} t_p^{\widetilde{\alpha}} = t_p^\alpha$.*

Moreover,

a) If $v_\alpha(t_p^\alpha) = p$, then $t_p^\alpha \leq \lim_{\widetilde{\alpha}\to\alpha^-} t_p^{\widetilde{\alpha}} \leq t^*$, with $t^* = \sup\{t : v_\alpha(t) = p\}$.

b) If $v_\alpha(t_p^\alpha) > p$, then $\lim_{\widetilde{\alpha}\to\alpha^-} t_p^{\widetilde{\alpha}} = t_p^\alpha$.

c) If $v_\alpha(\cdot)$ is an increasing and absolutely continuous function, then $\lim_{\widetilde{\alpha}\to\alpha} t_p^{\widetilde{\alpha}} = t_p^\alpha$.

2.4. Determining vaccination policies based on the quantiles of the extinction time. When an infectious disease is strongly detrimental for the population where it is spreading, such that it becomes an epidemic, then a vaccination policy should be applied to prevent the susceptible individuals and terminate the epidemic. Since it is impossible to immunize the whole population in most of the cases, only a proportion of susceptible individuals can be prevented by vaccination. How to determine this proportion is an important problem which depends on multiple factors. A significant factor for public authorities to assess the vaccination efficiency, is the time that the infectious disease should be allowed to survive after vaccination. To guarantee the extinction of the disease a.s., α should be at least equal to α_{inf}.

The vaccination policy determined by Gonzáles et al. (2010a) is based on the quantiles of the extinction time T_α. For fixed $0 < p < 1$ and $t > 0$, vaccination policies which guarantee that the infectious disease becomes extinct, with probability greater than or equal to p, not later than time t after the vaccination process ended, are constructed. Let us suppose that a proportion α of susceptible individuals are vaccinated. If there are z infected individuals at the end of the vaccination process, since each individual reproduces/infects independently from the others, then the probability that the disease becomes extinct no later than time t after vaccination process ended, can be bounded by $(v_\alpha(t))^z$.

Consequently, any vaccination level α such that $v_\alpha(t) \geq p^{(z)}$ or equivalently $t^\alpha_{p^{(z)}} \leq t$, with $p^{(z)} = p^{1/z}$, could be used. Taking this fact into account, as optimal vaccination policy is proposed that one, which corresponds to the smallest α of all of them, i.e.

$$\begin{aligned} \alpha_q = \alpha_q(p,t,z) &= \inf\{\alpha : \alpha_{\text{inf}} \leq \alpha \leq 1, v_\alpha(t) \geq p^{(z)}\} \\ &= \inf\{\alpha : \alpha_{\text{inf}} \leq \alpha \leq 1, t^\alpha_{p^{(z)}} \leq t\}. \end{aligned}$$

Applying the monotonicity and continuity properties of the functions $v_\alpha(t)$ and t^α_p (depending on α) we have that $v_{\alpha_q}(t) \geq p^{(z)}$ and $t^{\alpha_q}_{p^{(z)}} \leq t$ if $\alpha_q > \alpha_{\text{inf}}$.

3. Sevast'yanov's branching processes (SBP).

3.1. Properties of the time to extinction of the epidemic. For modelling the epidemic spread the probability that one infected individual with survival time (incubation plus contact periods) $u > 0$ contacts k healthy individuals, $k \geq 0$, is denoted by $p_k(u)$ and by α the proportion of immune individuals in the population. It is assumed that the population size is fixed and large enough so that α and the family of contact distribution laws, $\{p_k(u)\}_{k\geq 0}$, $u > 0$, can be considered stable along time. Then, the probability that an infected individual

with survival time $u > 0$ transmits the disease to k susceptible individuals is given by

$$p_{\alpha,k}(u) = \sum_{j=k}^{\infty} \binom{j}{k} \alpha^{j-k}(1-\alpha)^k p_j(u), \tag{7}$$

i.e., the infected individual with survival time u has been in contact with $j(= k, k+1, \ldots)$ healthy individuals and among them there have been k susceptible. The family $\{p_{\alpha,k}(u)\}_{k\geq 0}$, $u > 0$ is called, the infection distribution laws when the proportion of immune individuals in the population is α.

In parallel to the previous section the distribution of the time to extinction of a SBP depending on the vaccination level α with family of contact distribution laws $\{p_k(u)\}_{k\geq 0}$, $u > 0$ is investigated. To this end, for each α, by T_α is denoted the time to extinction of a SBP initiated at time 0 with a single infected individual, with family of infection distribution laws $\{p_{\alpha,k}(u)\}_{k\geq 0}$, $u > 0$, and with d.f. of the survival time $G(\cdot)$.

From now on, we denote by $v_\alpha(\cdot)$ the d.f. of the extinction time T_α, i.e. $v_\alpha(t) = P(T_\alpha \leq t)$ for all $t \in \mathbb{R}$. For each $u > 0$ we also denote by $f_\alpha(u, \cdot)$ the probability generating function (p.g.f.) of $\{p_{\alpha,k}(u)\}_{k\geq 0}$. Moreover, it is supposed that $G(0^+) = 0$, i.e., there is null probability of instantaneous *death* and consequently $v_\alpha(0) = 0$. Then, from Sevast'yanov (1971) we deduce that $v_\alpha(\cdot)$ is the unique bounded function such that

$$v_\alpha(t) = \begin{cases} 0, & t < 0, \\ \displaystyle\int_0^t f_\alpha(u, v_\alpha(t-u))dG(u), & t \geq 0. \end{cases} \tag{8}$$

Let $m = \displaystyle\int_0^\infty m(u)dG(u) < \infty$ and $m_\alpha = \displaystyle\int_0^\infty m_\alpha(u)dG(u) < \infty$, $0 \leq \alpha \leq 1$. Then, from (7) it is obtained that

$$m_\alpha = (1-\alpha)m. \tag{9}$$

Also, it is proved that

$$f_\alpha(u, s) = f(u, \alpha + (1-\alpha)s),\ 0 \leq s \leq 1,\ u > 0, \tag{10}$$

with $f(u, \cdot)$ the p.g.f. of the contact distribution law $\{p_k(u)\}_{k\geq 0}$, $u > 0$.

Let $q_\alpha = P(T_\alpha < \infty)$ be the extinction probability of a SBP with family of reproduction laws $\{p_{\alpha,k}(u)\}_{k\geq 0}$, $u > 0$. It is well known that $q_\alpha = 1$ iff $m_\alpha \leq 1$

(see Sevast'yanov (1971)). Notice that m_α is the critical threshold parameter of our model. So that, for such an α for which $m_\alpha > 1$, $v_\alpha(\cdot)$ is the d.f. of a non–proper random variable (r.v.) because $P(T_\alpha < \infty) < 1$.

From now on, it is considered those values of α, such that the extinction time T_α is a finite r.v., i.e. $m_\alpha \leq 1$, which implies that the infectious disease becomes extinct almost surely (a.s.). Taking into account (9), $m_\alpha \leq 1$ is equivalent to $\max\{0, 1 - m^{-1}\} \leq \alpha \leq 1$, which depends on the mean of contacts of an infected individual. In order to simplify the notations, by $\alpha_{inf} = \max\{0, 1 - m^{-1}\}$ is denoted the smallest proportion of immune individuals, so that the infectious disease becomes extinct a.s.

In González (2010b) the following results are proved:

Theorem 5. *If $0 \leq \alpha_1 < \alpha_2 \leq 1$, then $v_{\alpha_1}(t) \leq v_{\alpha_2}(t)$, for all $t \geq 0$.*

Theorem 6. *Let α be such that $m_\alpha < m_{\alpha_{inf}}$. Then for each $\varepsilon > 0$ there exists $\eta = \eta(\varepsilon, \alpha) > 0$ such that for all α^*, with $m_{\alpha^*} \leq 1$ and $|\alpha - \alpha^*| \leq \eta$,*

$$\sup_{0 \leq t < \infty} |v_\alpha(t) - v_{\alpha^*}(t)| \leq \varepsilon.$$

Furthermore, some parameters of T_α inherit these properties of $v_\alpha(\cdot)$. In what follows the monotonicity and the continuity properties of the mean of the distribution of the infection extinction time, depending on α are presented. Let's denote by μ_α the mean of time to extinction of infectious disease when the proportion of immune individuals is α. Since T_α is a non-negative r.v., then

$$\mu_\alpha = E[T_\alpha] = \int_0^\infty (1 - v_\alpha(t))dt. \tag{11}$$

Theorem 7.

1. *If $\alpha_{inf} \leq \alpha_1 < \alpha_2 \leq 1$, then $\mu_{\alpha_2} \leq \mu_{\alpha_1}$.*
2. *If $\overline{\alpha}$ is such that $0 < m_{\overline{\alpha}} < m_{\alpha_{inf}}$ and $\sup\{\mu_\alpha : \overline{\alpha} < \alpha \leq 1\} < \infty$, then $\mu_{\overline{\alpha}}$ is finite and $\mu_{\overline{\alpha}} = \lim_{\widetilde{\alpha} \to \overline{\alpha}^+} \mu_{\widetilde{\alpha}}$. Moreover, for all α with $\overline{\alpha} < \alpha \leq 1$, it follows that $\lim_{\widetilde{\alpha} \to \alpha} \mu_{\widetilde{\alpha}} = \mu_\alpha$.*

If the process starts with z infected individuals, then its time to extinction when the proportion of immune individuals in the population is α, will be $T_{\alpha,z} = \max\{T_\alpha^{(1)}, \ldots, T_\alpha^{(z)}\}$, where $T_\alpha^{(i)}$ are i.i.d. r.v. with the same distribution as T_α. So denoting by $v_{\alpha,z}(\cdot)$ the distribution function of $T_{\alpha,z}$, it follows that $v_{\alpha,z}(t) =$

$(v_\alpha(t))^z$, $t \in \mathbb{R}$. From this expression and considering the properties of the power functions, it is easy to establish for $v_{\alpha,z}(\cdot)$ the same monotonicity and continuity properties as those of $v_\alpha(\cdot)$. Moreover, these properties can be extended to the mean value of $T_{\alpha,z}$, denoted by $\mu_{\alpha,z}$.

3.2. Vaccination based on the mean value of the time to extinction. For fixed $\tau > 0$, the vaccination policies, which guarantee that the average time to extinction of an infection after vaccination period, t_1, is less than or equal to $t_1 + \tau$ are investigated.

The number of infected individuals at time t_1 is a random variable depending on α and on the number of infected individuals at the time t_0, when the vaccination process started. In the suggested modeling [20] it is approximated by its expected value. In general this is hard to calculate, but it is upper-bounded by the expected number of infected individuals at time t_1 providing the vaccination policy has not been applied.

Then, any vaccination level α such that $\mu_{\alpha,z} \leq \tau$ could be followed. The optimal vaccination policy is that one which corresponds to the smallest α, that is,

$$\alpha_{opt} = \alpha_{opt}(\tau, z) = \inf\{\alpha : \alpha_{\inf} \leq \alpha \leq 1, \mu_{\alpha,z} \leq \tau\}.$$

Taking into account the results of the previous section we have that $\mu_{\alpha_{opt},z} \leq \tau$ if $\alpha_{opt} > \alpha_{\inf}$. Therefore, vaccinating a proportion α_{opt} of susceptible individuals, the infectious disease becomes extinct in average, no latter than time τ after vaccination period. Moreover, although τ has been chosen arbitrarily, in order to find a solution of the problem, it is necessary that $\tau \geq \mu_{1,z}$.

3.3. Analyzing the control measures for avian influenza in Vietnam. The highly pathogenic H5N1 avian influenza virus has an incubation period after which it appears to be extremely virulent for a variety of domestic and wild bird species (see for example IDSA (2007)). The usual routes of bird-to-bird transmission are airborne transmission if birds are in close proximity, or direct contact with contaminated respiratory secretions. Also, since the contact period is considered to be very short (negligible) in comparison with the incubation period, an SBP is appropriate to model the spread of H5N1 virus in birds.

According to the official reports given by the World Organization for Animal Health (see the web page `http://www.oie.int`), Vietnam has been the country with greatest number of outbreaks of avian influenza in domestic birds from the end of 2003. In 7$^{\text{th}}$ December 2006 an outbreak started widespread itself in the southern part of the country and became extinct on 14$^{\text{th}}$ January 2007 (see OIE (2007)). The left plot of Figure 1 shows the numbers of infected domestic birds detected each day along this period. The non-null values are also given in Table 1.

Table 1. Non-null values of infected domestic birds detected between 7^{th} December 2006 and 14^{th} January 2007

Date	Cases	Date	Cases	Date	Cases	Date	Cases	Date	Cases
7 Dec	80	22 Dec	382	27 Dec	140	1 Jan	8	7 Jan	330
13 Dec	188	23 Dec	127	28 Dec	189	3 Jan	160	8 Jan	42
14 Dec	225	24 Dec	12	29 Dec	60	4 Jan	378	9 Jan	10
19 Dec	6073	25 Dec	262	30 Dec	18	5 Jan	240	12 Jan	880
20 Dec	40	26 Dec	1908	31 Dec	130	6 Jan	300	14 Jan	1621

From 20^{th} December the number of cases decreases, probably because some control measures were taken (see OIE (2007)).

Next, the spread of the H5N1 avian influenza virus in Vietnam from 19^{th} December until 14^{th} January is analyzed by comparing it with the simulated times to extinction of SBP for different vaccination levels. González et al. (2010b) have considered that $G(\cdot)$ is gamma d.f. and, for each $u > 0$, $\{p_k(u)\}_{k\geq 0}$ follows a Poisson distribution with parameter λu, being $\lambda > 0$. These types of distributions have been found to be appropriate for the survival time (including incubation and contact periods) and the number of contacts, respectively (see Daley et al. (1999), Farrington et al. (1999, 2003b) or Mode et al. (2000)).

Taking into account that the incubation period of H5N1 avian influenza virus is estimated between 3 and 7 days (see IDSA (2003)) the gamma distribution with mean 5 and shape 16 is considered, to guarantee that the survival period in 90% of individuals is between 3 and 7 days. Therefore, it is deduced that $m = 5\lambda$. Since the number of infected individuals at the first outbreak (on 7^{th} December) is 80, and after the incubation period (in 13^{th} and 14^{th} December) the total number of infected individuals was 413, the rate m, using Lotka's estimator, can be estimated as $\widehat{m} = 2132$ (see Guttorp (1991)). No more outbreaks were taken into account, as according to what is observed above, some control measures have been applied before 19^{th} December. Thus, in order to apply the method, González et al. (2010b) consider this date as the end of vaccination period. The number of individuals incubating the virus at this date is estimated at $\widehat{m} \simeq 2132$. Finally, for each vaccination level, α, $0 \leq \alpha \leq 1$, it is deduced from (10) that $\{p_{k,\alpha}(u)\}_{k\geq 0}$ also follows a Poisson distribution with parameter $u(1-\alpha)\lambda$, $u > 0$.

The right–hand plot of Figure 1 shows the histogram of 10,000 simulated times to extinction for $\alpha = 1$. Assuming that the model fits well, from the fact that the virus took close to 30 days to become extinct after the vaccination period ended, while the maximum of simulated extinction times is less than 30, it is deduced

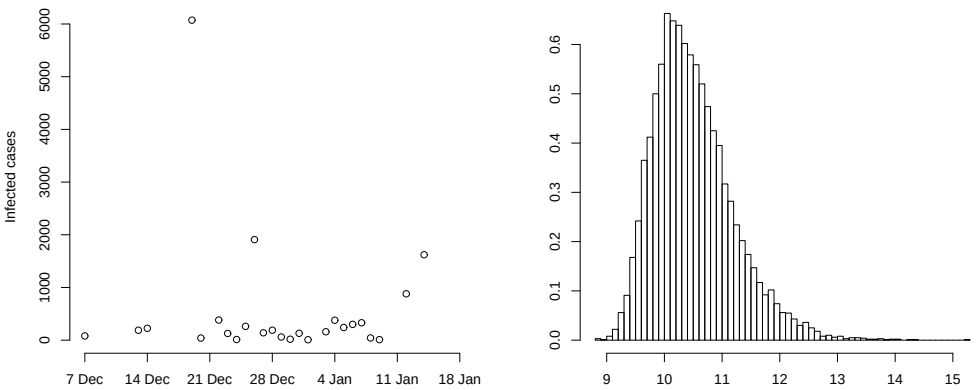

Fig. 1. Left: Numbers of infected domestic birds detected between 7^{th} December 2006 and 14^{th} January 2007. Right: Histogram of simulated times to extinction for $\alpha = 1$

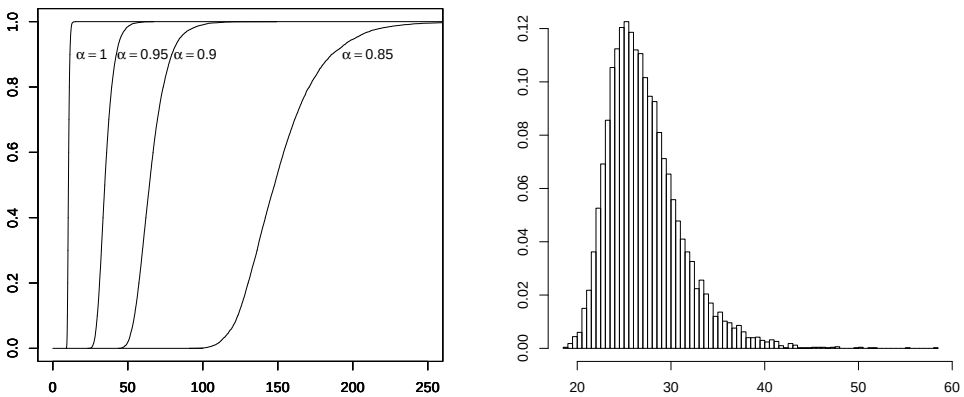

Fig. 2. Left: Empirical d.f. of the time to extinction for $\alpha = 0.85, 0.90, 0.95$ and 1. Right: Histogram of simulated extinction times for $\alpha = 0.97$

that the control measures followed in Vietnam did not cover all the susceptible individuals. Consequently, the control measures in Vietnam correspond to a vaccination level $\alpha < 1$ in the above setting. Let us now determine α_{opt} which corresponds to these control measures. From Theorem 5 it is deduced that the smaller is α the longer the time to extinction. This behaviour is shown in the left–hand panel of Figure 2 where the empirical d.f. of the time to extinction is plotted for $\alpha = 1, 0.95, 0.90$ and 0.85. Since the virus took close to 30 days to become extinct, then the vaccination level must have been close to 1. Taking into account the vaccination policy based on the mean value of the time to extinction, by applying the simulation–based method, it is obtained that $\alpha_{opt}(\tau = 30, z = 2132) = 0.97$. The right–hand panel of Figure 2 shows the histogram of 10,000 simulated times to extinction for $\alpha = 0.97$. In conclusion, the control strategies followed in Vietnam correspond to a vaccination level close to 1 ($\alpha_{opt} = 0.97$). Of course one must have in mind that such a high proportion is connected with the high risk of death not only in the birds but also in the human population in the case of of bird–to–human transmission.

4. Bayesian approach for predicting outbreaks. Bayesian approach for predicting outbreaks, implemented in the statistical software R (see Höhle (2005)) is applied on surveillance data of mumps collected in Bulgaria for the period 2000–2008. A detailed description of the method could be seen in Höhle (2005). The official data is kindly provided by the National Center of Infectious and Parasitic diseases at the Ministry of Health, Bulgaria. It has been collected on a weekly base and presents the epidemic picture by regions in the country for 2000–2008 year. The data clearly shows that there was epidemic outbreak in the country in 2007 and 2008 (see Kojouharova et al. (2007)).

Using R–software in Mitova–Bobcheva et al. (2011) the following model of the data was applied. Let us denote by $\{y_t;\ t = 1, \ldots, n\}$ the time series of counts representing the surveillance data. Due to the fact that such data is typically collected on a weekly basis it is also convenient to use the following notation $\{y_{i:j}\}$, where $j = \{1, \ldots, 52\}$ presents the number of weeks in the year and $i = \{-b, \ldots, -1, 0\}$ are the corresponding years. The years have been indexed in such a way that the last year for which we have data has index 0. Let $y_{0:t}$ be the number of cases of the current week (denoted week t in year 0), b the number of years to go back in time and w the number of years around t to include from those previous years. The zero year will be denoted by w_0. Hence the set of chosen weeks/years for which we want to trace the disease is:

$$R(w, w_0, b) = \left(\bigcup_{i=1}^{b} \bigcup_{j=-w}^{w} y_{-i:t+j} \right) \cup \left(\bigcup_{k=-w_0}^{-1} y_{0:t+k} \right).$$

Note that the number of cases of the current week is not a part of $R(w, w_0, b)$. The aim of the surveillance algorithm described above is to create a prediction $\hat{y_{t:0}}$ for the current week of the process. This prediction is then compared to the actual observed value $y_{0:t}$. If the observed value is much higher than the predicted one we get an alert, which warns us to investigate further the reasons for this.

More applications and details one could see in Mitova–Bobcheva et al. (2011).

5. Bayesian estimation of the offspring mean.

5.1. Biological background and motivation. The fundamental epidemiological quantity determining whether an infectious disease will persist in a host population is the basic reproduction number, R_0 (Anderson et al. (1991) and Heesterbeek et al. (1996)). This is defined as the average number of secondary infections caused in a susceptible population by a typical infected. R_0 is a key factor in determining how fast an infection will spread in a population. If $R_0 > 1$, the infectious agent has the potential to persist indefinitely, whilst if $R_0 < 1$, the incidence of infection will decay to zero. The reason is clear: if a primary infection is unable to generate at least one replacement secondary infection, the numbers of infected in the population will inevitably decline through time.

The work by Angelov et al. presents a Bayesian approach of estimating R_0 for infectious diseases like mumps, measles and possibly others, that follows so-called SIR (susceptible → infective → removed) and SEIR (susceptible → exposed → infective → removed) scheme in epidemiological context, from the case data comprising of the number of infected on a weekly base.

Under the assumption that each infective infects a random number of individuals in accordance with some probability distribution and that this distribution does not change over time and is the same for all individuals, it is reasonable to model the number of infected by a branching process. The simplest class of branching processes – Bienaymé–Galton–Watson processes is used. In fact, the assumption that the distribution of the number of infected individuals by one infectious does not change over time, is not always realistic, because increasing the number of infectious individuals reduces the number of susceptible to the disease. However, in populations with large number of susceptible – over 100, this assumption is not away from reality (see Farrington et al. (2003)). Since these are discrete time processes, the number of infected by each infectious is not

counted in real time, but at the end of its infectious period (the period during which one infective could transmit the disease to others susceptible). Despite its idealization, such models are widely used in epidemiology, for example see Becker (1974), Heyde (1979), Farrington et al. (1999, 2003), Yanev et al. (1999). More complex branching process also have been applied for modeling of infectious diseases, see Marschner (1992), Ball et al. (1995), Becker et al. (2004), González et al. (2009, 2010b) and Jacob (2010).

Usually the information about the spread of the disease is not complete – do not know the number of infected by each infectious individual. Models of branching processes and application of Bayesian methods allow to estimate the basic reproduction number R_0 using data on reported cases, collected by institutions for control of public health. A similar approach was proposed by Farrington et al. (2003b).

The statistical inference is applied to real data on the number of reported cases of mumps in Bulgaria during the period 2005–2008 provided by the National Center of Infectious and Parasitic Diseases. It is assumed that the offspring distribution of the branching process belongs to the family of generalized power series distributions, which is quite a broad class of discrete distributions, including binomial, Poisson and geometric ones. It turned out that for this wide class of distributions, it is possible to obtain exactly the distribution of the total progeny of the BGWBP, which is needed for estimation of offspring mean λ. Point and interval estimates of λ, applying a Bayesian approach by simulating the posterior distribution using Metropolis–Hastings algorithm are found. The algorithm is implemented in the language and environment for statistical computing `R`, version 2.11.1 (see R Development team R Development Core Team (2010)).

5.2. Bienaymé–Galton–Watson BP with power series offspring distribution as a model of epidemic spread From now on it is assumed that X has a generalized power series distribution, i. e.

$$P(X = k) = \frac{a_k \theta^k}{A(\theta)}, \quad k \in \mathcal{K}$$

where $a_k \geq 0$, $A(\theta) = \sum_k a_k \theta^k$, $\theta > 0$, $\mathcal{K} \subseteq \{0, 1, 2, \ldots\}$. The parameter θ is called canonical parameter. Distributions of this type are the binomial, Poisson, negative binomial (in particular – the geometric). The mean of X is

$$\lambda = EX = \frac{\theta A'(\theta)}{A(\theta)}.$$

As it is noticed, one of the reasons to use branching processes as models of infectious disease spread is the obvious fact, that the offspring mean λ is identified as a basic reproduction number R_0 in epidemiology. The task is to estimate λ on the basis of data on the number of infected individuals. Most often the data on the number of infected ones by each infectious is missing, but what is available, in fact is of the total number of infected individuals for a given period of time. Therefore, the estimation of λ is based on the total number of infected individuals by the end of the outbreak, called a total progeny in a branching processes' context.

Let us denote by Y the total progeny of BGWBP or the total number of infected individuals by the end of the outbreak. It is defined as follows

$$Y = \sum_{n=0}^{\infty} Z_n.$$

Then as a consequence, the distribution of Y has the form

$$P(Y = r) = \frac{s}{r} P(X_1 + X_2 + \ldots + X_r = r - s), \qquad r = s,\ s+1,\ s+2, \ldots$$

where $X_1, X_2, \ldots, X_r$ are i.i.d.r.v. with the same distribution as X (see Jagers (1975)). It is obvious that the distribution of Y is given by r^{th} convolution of X.

In what follows the method of obtaining total progeny distribution given the offspring one, in particular cases of Poisson and geometric offspring distributions is shortly presented (for more details see Angelov et al. (2012)). Geometric and Poisson offspring distributions correspond respectively to the limiting branching process for a general stochastic epidemic and a Reed-Frost epidemic model (see Ball (1983)).

Poisson offspring. Let the offspring distribution be Poisson:

$$P(X = k) = \frac{e^{-\lambda}\, \lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

Using that the sum of r i.i.d. Poisson r.v. has Poisson distribution with parameter λr we directly express:

$$P(X_1 + X_2 + \ldots + X_r = k) = \frac{e^{-\lambda r} (\lambda r)^k}{k!}.$$

Thus the distribution of the total progeny is:

$$P(Y = r) \;\; = \;\; \frac{s}{r} P(X_1 + X_2 + \ldots + X_r = r - s)$$

$$= \frac{s}{r}\frac{e^{-\lambda r}(\lambda r)^{r-s}}{(r-s)!}, \qquad r = s,\ s+1,\ s+2, \ldots,$$

i.e. Y has a Borel–Tanner distribution (see Haight et al. (1960)).

In an analogous way it is obtained (see Angelov et al. (2012)) that for geometric offspring distribution Y has a distribution of Haight (see Haight (1961)).

5.3. Bayesian estimation of λ. The Metropolis–Hastings algorithm is used, with which some computational difficulties in Bayesian estimation could be avoided. More details on this topic can be found in Robert (2007), Robert et al. (2004, 2010) and Hoff (2009).

Actually, λ is estimated having data from a single outbreak, i.e. knowing that the total number of infected is y, and the initial number of infected is s. In this case the likelihood function for λ has the form:

$$L(y|\lambda) = P(Y = y; s, \lambda).$$

Following a Bayesian approach, it is assumed that the parameter λ is a random variable with prior distribution $\pi(\lambda)$. Then the posterior density is given by the Bayes' formula:

$$f(\lambda|y) = \frac{L(y|\lambda)\pi(\lambda)}{\int_0^\infty L(y|\lambda)\pi(\lambda)d\lambda}.$$

Using squared error loss function, the Bayesian estimate of λ, will be the mean of the posterior distribution:

$$\widehat{\lambda} = E(\lambda|y).$$

Concerning the interval estimation of λ, let us recall that the interval $[a, b]$ is called $100(1-\alpha)\%$ highest posterior density interval (HPDI) for parameter λ, if the following conditions are satisfied:

(a1) $P(\lambda \in [a, b] \mid y) = 1 - \alpha$, for a fixed $\alpha \in (0, 1)$;

(a2) If $\lambda_1 \in [a, b]$ and $\lambda_2 \notin [a, b]$, then $f(\lambda_1|y) > f(\lambda_2|y)$.

In general, the explicit calculation of the posterior density $f(\lambda|y)$ is difficult. To avoid such difficulties, Metropolis–Hastings sampling based on random walk to evaluate the posterior distribution, is used. This algorithm allows to

simulate any random variable, if its density is known up to a normalizing constant, in above case: $f(\lambda|y) = cL(y|\lambda)\pi(\lambda)$ and it is not necessary to calculate $c = 1/\int_0^\infty L(y|\lambda)\pi(\lambda)d\lambda$.

After generating $\lambda_1, \lambda_2, \ldots, \lambda_N \sim f(\lambda|y)$ their empirical distribution is used as an approximation of $f(\lambda|y)$. So the Bayesian estimate of λ is:

$$\widehat{\lambda} = \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_N}{N}.$$

As prior distributions for λ will be considered uniform $U[0, 2]$ and log-normal $LN(\mu = 0, \sigma = 1)$. Both have median 1, i.e., are neutral with respect to whether $\lambda < 1$ or $\lambda > 1$.

Angelov et al. (2012) consider two cases for offspring distribution – Poisson and geometric, the likelihood function $L(y|\lambda)$ to be the Borel–Tanner probability mass function and the Haight probability mass function, respectively.

5.4. Mumps in Bulgaria – estimation of reproduction number. In this subsection the described methods for estimation of offspring mean of BG-WBP is illustrated, using data on the number of reported cases of mumps in Bulgaria during the period 2005–2008.

Mumps is a viral infectious disease of humans and spreads from person to person through the air. The period between mumps transmission and the beginning of mumps symptoms is called the incubation period for mumps. This period is between 14 and 24 days (median 18 days). The infectious period starts about 2 days before the onset of symptoms and usually, an individual with mumps symptoms is immediately isolated from the population. In view of the length of the incubation period, it is considered that an outbreak in a region is a sequence of weeks with no more than three consecutive weeks without cases. That is, when more than three weeks without cases are observed, the outbreak is considered that has become extinct, with the next outbreak starting in the first subsequent week in which there is at least one new case.

In 2007 in Bulgaria there was an outbreak of mumps. Over 60% of those infected at the beginning of the year are aged between 15 and 19 years, about 20% between 20 and 24 years. It is assumed that the outbreak was the result of poor immunization policy in the 80s. One third of patients aged between 15 and 19 years have never been vaccinated, about half was given only one dose of vaccine, which is found not effective. Over 90% of 20–24–years–old have not been vaccinated against mumps (see Kojouharova (2007)).

The data, provided by the National Center of Infectious and Parasitic Diseases, consists of the number of reported cases of mumps in Bulgaria during the

period 2005 to 2008, on weekly base for each of 28 regions of the country. The 28 regions are treated separately.

Each outbreak is considered as a realization of a branching process. The data that is observed about the process are the total number y of infected and the initial s number of infectious. The reproduction number for the outbreaks in Sofia–city is estimated (for the regions of Kyustendil and Lovech one can see Angelov et al. (2012)). For the offspring distribution are used – Poisson and geometric ones and for each of them two prior distributions – uniform and log-normal are suggested, so a total of four estimates for λ for each region are obtained. For each of the options 5000 random numbers are generated with the corresponding posterior distribution and the first 500 are ignored. For calculating highest posterior density interval we use the function `HPDinterval` from `coda` package (see Plummer et al. (2010)).

In Sofia–city during the period from the 40^{th} week of 2006 to the 52^{nd} week of 2008 a total number of 2124 cases of mumps was reported and the initial number of infectious individuals was 2, i.e. $y = 2124; s = 2$. Point estimates for λ and HPD intervals (95% HPDI = 95 percent highest posterior density interval) are given in Table 2.

Table 2. Point and interval estimates of λ for Sofia–city

	Offspring distribution	*Prior distribution*	$\widehat{\lambda}$	95% HPDI
1	Poisson	Uniform	1.0011	$[0.9577, 1.0436]$
2	Poisson	Log–normal	0.9981	$[0.9540, 1.0412]$
3	Geometric	Uniform	1.0002	$[0.9459, 1.0646]$
4	Geometric	Log–normal	0.9996	$[0.9383, 1.0598]$

One can see that the estimates $\widehat{\lambda}$ and HPD intervals are quite close for different assumptions about offspring and prior distributions.

All developed epidemic models by means of continuous–time BP are aimed to take into account the variability of many factors – infection, incubation period and survival, and incorporating one specific characteristic of BP, namely the extinction time, make it possible to develop different scenarios for additional vaccination in order to prevent the population from most dangerous evolution.

REFERENCES

[1] A. Agresti. Bounds on the extinction time distribution of a branching process. *Adv. Appl. Probab.* **6** (1974), 332–335.

[2] H. Andersson, T. Britton. Stochastic epidemic models and their statistical analysis. Lecture Notes in Statistics, vol. **151** (2000), Springer–Verlag.

[3] R. M. Anderson, R. M. May. Infectious Diseases of Humans: Dynamics and Control, Oxford University Press, 1991.

[4] A. Angelov, M. Slavtchova–Bojkova. Bayesian estimation of the offspring mean in branching processes: Application to infectious disease data. *Computers and Mathematics with Applications*, (2012) doi: 10.1016/j.camwa.2012.01.049.

[5] K. Athreya, P. Ney. Branching Processes. Berlin, Springer–Verlag, 1972.

[6] D. E. Axelrod, Y. Gusev, J. W. Gamel. Ras-oncogene transformed and non–transformed cell population are each heterogeneous but respond differently to the chemotherapeutic drug cytosine arabinoside (Ara–C). *Cancer Chemotherapy and Pharmacology* **29** (1997), 445–451.

[7] D. E. Axelrod, Y. Gusev, T. Kuczek. Persistence of cell cycle times over many generations as determined by heriatability of colony sizes of ras oncogene-transformed and non-transformed cells. *Cell Proliferation* **26** (1993), 235–249.

[8] F. Ball. The threshold behaviour of epidemic models. *J. Appl. Probab.* **20** (1983), 227–241.

[9] F. Ball, P. Donnelley. Strong approximations for epidemic models. *Stochastic Process. Appl.* **55** (1995), 1–21.

[10] A. D. Barbour. The duration of the closed stochastic epidemic. *Biometrika* **62** (1975), 477–482.

[11] N. Becker. On parametric estimation for mortal branching processes. *Biometrika* **61** (1974), 393–399.

[12] N. Becker, T. Britton. Estimating vaccine efficacy from small outbreaks. *Biometrika* **91** (2004), 363–382.

[13] D. J. Daley, J. Gani. Epidemic modelling: An introduction. Cambridge University Press, 1999.

[14] C. Farrington, N. Andrews. Monitoring the Health of Populations. Chapter Outbreak Detection: Application to Infectious Disease Surveillance, Oxford University Press, 2003a.

[15] C. Farrington, N. Andrews, A. Beale, M. Catchpole. A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society* Series A. **159** (1996), 547–563.

[16] C. Farrington, A. Grant. The distribution of time to extinction in subcritical branching processes: applications to outbreaks of infectious disease. *J. Appl. Probab.* **36** (1999), 771–779.

[17] C. P. Farrington, M. N. Kanaan, N. J. Gay. Branching process models for surveillance of infectious diseases controlled by mass vaccination. *Biostatistics* **4** (2003b), No 2, 279–295.

[18] M. González, R. Martínez, M. Slavtchova–Bojkova. Age–dependent branching processes as models of infectious diseases. *C. R. Acad. Bulg. Sci.* **62** (2009), No 5, 541–550.

[19] M. González, R. Martínez, M. Slavtchova–Bojkova. Stochastic monotonicity and continuity properties of the extinction time of Bellman–Harris branching processes: an application to epidemic modelling. *Jour. Appl. Prob.* **47** (2010a), 58–71.

[20] M. González, R. Martínez, M. Slavtchova–Bojkova. Time to extinction of infectious diseases through age-dependent branching models. *Lecture Notes in Statistics–Proceedings* **197** (2010b), 241–256.

[21] P. Guttorp. Statistical Inference for Branching Processes. John Wiley and Sons, Inc., 1991.

[22] P. Haccou, P. Jagers, V. Vatutin. Branching processes: variation, growth and extinction of populations. Cambridge University Press, 2005.

[23] F. A. Haight. A distribution analogous to the Borel–Tanner. *Biometrika* **48** (1961), 167–173.

[24] F. A. Haight, M. A. Breuer. The Borel–Tanner distribution. *Biometrika* **47** (1960), 143–150.

[25] J. A. P. Heesterbeek, K. Dietz. The concept of R_0 in epidemic theory. *Statistica Nederlandica* **50** (1996), 89–110.

[26] D. Heinzmann. Extinction times in multitype Markov branching processes. *J. Appl. Probab.* **46** (2009), No 1, 296–307.

[27] C. C. Heyde. On assessing the potential severity of an outbreak of a rare infectious disease: a Bayesian approach. *Australian Journal of Statistics* **21** (1979), 282–292.

[28] P.D. Hoff. A First Course in Bayesian Statistical Methods. Springer, 2009.

[29] R. Höhle. The R–package "surveillance". Sonderforschungsbereich 386, 2005.

[30] IDSA: Avian influenza (bird flu): agricultural and wildlife considerations. Infectious Diseases Society of America, 2007, `http://www.idsociety.org/pandemicinfluenza.htm`.

[31] V. Isham. Stochastic models for epidemics. Celebrating statistics papers in honour of Sir David Cox on his 80th birthday (Eds A. C. Davison, Y. Dodge, N. Wermuth) Chapter 1, Oxford University Press, 2005.

[32] C. Jacob. Branching processes: their role in epidemiology, *International Journal of Environmental Research and Public Health* **7** (2010), 1186–1204.

[33] P. Jagers. Branching Processes with Biological Applications. John Wiley and Sons, 1975.

[34] M. Kimmel. Nonparametric analysis of stathmokinesis. *Math. Biosci.* **74** (1985), 111–123.

[35] M. Kimmel, D. Axelrod. Branching processes in Biology. Springer–Verlag, 2002.

[36] M. Kimmel, F. Traganos. Estimation and prediction of cell cycle speciffic effects of anticancer drugs. *Math. Biosci* **80** (1986), 187–208.

[37] M. Kojouharova, A. Kurchatova, L. Marinova, T. Georgieva. Mumps outbreak in Bulgaria, 2007: a preliminary report. *Eurosurveillance*, **12** (2007).

[38] I. C. Marschner. The effect of preferential mixing on the growth of an epidemic. *Mathematical Biosciences* **109** (1992), 39–67.

[39] M. Mitova–Bobcheva, M. Slavtchova–Bojkova, M. Kojouharova, A. Kurchatova. Analysing and monitoring surveillance data of munmps in Bulgaria. *Pliska Stud. Math. Bulgar.* **20** (2011), 149-–154.

[40] C. J. Mode, C. K. Sleeman. Stochastic processes in epidemiology. World Scientific, 2000.

[41] I. Nasell. Stochastic models of some endemic infections. *Math. Biosci.* **179** (2002), 1–19.

[42] OIE: Report reference 1828/TY–DT. World Organization for Animal Health, 2007. `http://www.oie.int`.

[43] A. Pakes. On the asymptotic behaviour of the extinction time of the simple branching process. *Adv. Appl. Probab.* **21** (1989), 470–472.

[44] A. Pakes. Biological applications of branching processes. Handbook of Statistics Vol. 21 Stochastic Processes: Modelling and Simulation (Eds D. N. Shanbhag, C. R. Rao) Chapter 18, 693–773, Elsevier Science B.V., 2003.

[45] M. Plummer, N. Best, K. Cowles, K. Vines. coda: Output analysis and diagnostics for MCMC. R package version 0.13-5. `http://CRAN.R-project.org/package=coda`, 2010.

[46] R Development Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria. ISBN 3-900051-07-0, 2010.

[47] C. Robert. The Bayesian Choice, 2nd edition, Springer, 2007.

[48] C. Robert, G. Casella. Monte Carlo Statistical Methods, 2nd edition, Springer, 2004.

[49] C. Robert, G. Casella. Introducing Monte Carlo Methods with R, Springer, 2010.

[50] B. A. Sevast'yanov. Branching Processes, Mir, Moscow, 1971 (in Russian).

[51] M. Slavtchova–Bojkova, N. M. Yanev. Branching Stochastic Processes, Sofia University Press "St. Kliment Ohridski", ISBN 978-954-07-2601-4, 2007 (in Bulgarian).

[52] D. Stroup, G. Williamson, J. Herndon, J. Karon. Detection of aberrations in the occurrence of notifiable diseases surveillance data. *Statistics in Medicine* **8** (1989), 323–329.

[53] A. Yakovlev, N. Yanev. Branching stochastic processes with immigration in analysis of renewing cell populations. *Math. Biosci.* **203** (2006), 37–63.

[54] A. Yakovlev, N. Yanev. Age and residual lifetime distributions for branching processes. *Statist. Probab. Lett.* **77** (2007), No 5, 503–513.

[55] G. P. Yanev, C. P. Tsokos. Decision–theoretic estimation of the offspring mean in mortal branching processes, *Communications in Statistics: Stochastic Models* **15** (1999), 889–902.

Maroussia Nikiforova Slavtchova–Bojkova
Department of Probability, Operations Research and Statistics
Faculty of Mathematics and Informatics
St. Kl. Ohridski University of Sofia
5, J. Bourchier Blvd
1164 Sofia, Bulgaria
e-mail: `bojkova@fmi.uni-sofia.bg`
and
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Acad. G. Bonchev Str., Bl. 8
1113 Sofia, Bulgaria